Cloning and characterization of the human osteopontin gene and its promoter

Naoki HIJIYA, Mihoko SETOGUCHI, Keiko MATSUURA, Yasunori HIGUCHI, Shin'ichiro AKIZUKI and Shunsuke YAMAMOTO*
Department of Pathology, Oita Medical University, Hasama-machi, Oita 879-55, Japan

We isolated the human osteopontin (hOP) gene and the 5′ upstream region, and analysed its exon–intron structure and potential regulatory sequences of the promoter region in comparison with those of the mouse and porcine gene. The coding sequence is split into 7 exons which are similar to those of the mouse gene, although the hOP gene is longer than the mouse gene. The difference in length is mainly due to variations in intron 3, which is ~ 2.7-fold longer than that of the mouse OP gene. The 5′ upstream region of the hOP, which is highly conserved up to nucleotide −250, contains a number of potential *cis* regulatory consensus sequences. A series of sequentially 5′-deleted chimeric clones was tested for the ability to stimulate chloramphenicol acetyltransferase (CAT). Initial CAT analysis demonstrated that nucleotides at positions −474 to −270, −124 to −80, and −55 to −39 contained *cis*-acting enhancing sequences in a human monocyte cell line, SCC-3, although the −124 to −80 region was much more active than other regions. Deletion of the sequences between −474 and −270 localized this *cis* region to the sequence at positions −439 to −410, whereas the deletion between −124 to −80 localized the regions to −124 to −115, and −94 to −80. Gel-shift analysis using as probes synthesized double-stranded DNA corresponding to the 10 and 15 bp region at positions −124 to −115 and −94 to −80 respectively revealed that each probe formed a major band complexed with nuclear proteins prepared from SCC-3 cells.

INTRODUCTION

Osteopontin (OP) is a sialoprotein of ~ 298 amino acids which contains a Gly-Arg-Gly-Asp-Ser (GRGDS) sequence and is a highly hydrophilic and negatively charged structure, consisting mainly of successive aspartic acids in the central region of the molecule [1–5]. The roles of this uniquely structured protein in cell adhesion and mineral binding have been reported [1,6–11].

Although expression of OP is not ubiquitous, OP is expressed in a variety of cells, and shows diverse features of expression. Constitutive expression of OP is found in bone, kidney, placenta, nerve cells and macrophages, whereas inducible or enhanced expression of OP is observed in T lymphocytes [12], epidermal and bone cells and macrophages with a variety of agents; these include the tumour promoter phorbol 12-myristate 13-acetate (PMA), 1,25-dihydroxyvitamin D_3, basic fibroblast growth factor, leukaemia inhibitory factor, tumour necrosis factor-α, interleukin-1 [1,13–21], lipopolysaccharide and interferon-γ. Furthermore, OP is expressed in the neoplastic state [22]. These results suggest that the expression of OP is controlled by complex regulatory systems, which could differ among cells.

We have previously cloned the mouse OP (mOP) cDNA and gene, and deduced the amino acid sequence and structure of the mOP gene [2,16]. We reported that the structure consists of six exons, although another exon further upstream has been demonstrated more recently [23,24]. A similar gene organization consisting of seven exons was also suggested for the porcine OP gene [25]. In this paper, we present the complete structure of the human OP (hOP) gene and its 5′ upstream sequence, and compare its organization and potential regulatory sequences with those of the mouse and pig. In addition, we have analysed regulatory elements for the constitutive expression of the hOP gene in the human monocyte cell line SCC-3, which expresses strongly the hOP gene.

EXPERIMENTAL

Materials

Restriction enzymes were purchased from Takara Shuzo Co. (Kyoto, Japan), Toyobo Co. (Osaka, Japan), Bethesda Research Laboratories (Gaithersburg, MD, U.S.A.), and Wako Pure Chemicals Co. (Tokyo, Japan). Agarose, ultrapure DNA grade, and the DNA-ligation kit were from Takara. A 5′-deletion kit was obtained from Nippongene Co. (Tokyo, Japan). The radioactive nucleotides [α-^{32}P]dCTP (3000 Ci/mmol) and [γ-^{32}P]ATP (6000 Ci/mmol) were obtained from Du Pont–New England Nuclear. Chemicals used for DNA sequencing were obtained from Toyobo Co. X-ray film (XAR-351) was obtained from Kodak.

Cell lines and cell cultures

The human monocyte cell lines SCC-3 [26] and THP-1 [27], the human B lymphoma cell line Raji [28], and the mouse macrophage cell line HINS-B3 [29] were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% (v/v) fetal calf serum.

Molecular cloning and sequencing

The hOP gene and the 5′ upstream sequence up to −4 kb were cloned from a bacteriophage λ GEM-11 human liver library (Promega) using ^{32}P-labelled cDNA containing the hOP-1 insert, which was isolated from a human kidney cDNA library

Abbreviations used: DTT, dithiothreitol; DMEM, Dulbecco's modified Eagle's medium; mOP, mouse OP; OP, osteopontin; PMA, phorbol 12-myristate 13-acetate; PMSF, phenylmethanesulphonyl fluoride; pOP, porcine OP; VDR, vitamin-D-responsive sequence; hOP, human osteopontin; CAT, chloramphenicol acetyltransferase.

* To whom correspondence should be addressed.

The nucleotide sequences reported in this paper have been submitted to the DDBJ/EMBL/GenBank Data Bank with accession numbers D14813 (the human osteopontin gene) and D14816 (5′ upstream of the mouse osteopontin gene).

(Clontech) using mOP cDNA [2] and contained sequence from nucleotides position 177 to 878 [4]. Sequencing was performed by the dideoxy-chain-termination method of Sanger et al. [30]. The DNA sequence was determined by isolation and subcloning of specific restriction fragments into either pUC118 or pUC119 vector. The entire nucleotide sequence of the 5′ upstream, exon and intron was determined from both strands.

S1 nuclease mapping

S1 nuclease mapping was carried out according to the method of Berk and Sharp [31]. A single-stranded DNA corresponding to the *Xba*I–*Xba*I (nucleotides −581 to +260, 841 bp) fragment containing the region from the 5′ upstream to intron 1 of the hOP gene was prepared using M13K07 phage as a vector. An oligonucleotide, 5′-CTTGGTCGGCGTTTGGCTGAGAAG-GCTGCA-3′, which is complementary to positions +61 to +90 of the first exon of the hOP gene, was labelled at its 5′ terminus with [γ-^{32}P]ATP and T4 polynucleotide kinase. The *Xba*I–*Xba*I fragment was annealed with the end-labelled oligonucleotide primer. The annealed DNA was incubated with 400 μM dNTPs and 10 units of the Klenow fragment of *Escherichia coli* DNA polymerase for 30 min at 37 °C. The mixture was heated to 65 °C for 5 min to inactivate the Klenow fragment, then chilled on ice. After digestion with *Bam*HI and alkali denaturation, the extended primer was separated on alkali gel electrophoresis.

Total cellular RNA (50 μg from THP-1 cells) was hybridized for 16 h at 30 °C with the extended primer in 40 mM Pipes, pH 6.4, 0.4 M NaCl, 1 mM EDTA and 80% (v/v) formamide. Following hybridization, the reaction was diluted 10-fold with S1 nuclease buffer (280 mM NaCl, 50 mM sodium acetate, pH 4.5, and 4.5 mM $ZnCl_2$) and 6 μg of salmon sperm DNA. S1 nuclease (300 units) was added and the reaction mixture was incubated for 1 h at 37 °C. The reaction mixture was terminated by addition of termination buffer (4 M ammonium acetate, 20 mM EDTA), and the DNA–RNA hybrids were precipitated with ethanol, dissolved in the loading buffer, heated to 90 °C, and resolved on an 6% (w/v) acrylamide/8 M urea sequencing gel.

Primer extension analysis

The probe for primer extension analysis was a synthetic 30 base single-stranded end-labelled primer used in the S1 nuclease protection. The primer was annealed to the 50 μg of total cellular RNA from THP-1 cells by heating the reaction mixture for 1 h at 80 °C in 20 μl of 80% formamide, containing 400 mM NaCl, 40 mM Pipes (pH 6.4) and 1 mM EDTA, followed by incubation for 12 h at 30 °C. The resulting DNA–RNA hybrid was precipitated in ethanol and dissolved in reverse transcriptase buffer (50 mM Tris/HCl, pH 8.0, 20 mM 2-mercaptoethanol, 100 mM KCl, 10 mM $MgCl_2$), 1 mM deoxynucleotides, 40 units of reverse transcriptase and 60 units of RNAase inhibitor. After 1.5 h at 42 °C, the DNA–RNA hybrids were incubated with 1 μl of 0.5 M EDTA and 1 μg of RNAase for 10 min at 37 °C, then phenol-extracted, ethanol-precipitated, dissolved in loading buffer [95% formamide, 20 mM EDTA, 0.05% (v/v) Bromophenol Blue, 0.05% (v/v) Xylene Cyanol FF], heated to 90 °C, and resolved on a 6% acrylamide/8 M urea sequencing gel.

Construction of hOP–chloramphenicol acetyltransferase (CAT) plasmids

*Sac*I–*Sty*I, *Eco*RI–*Sty*I, *Xba*I–*Sty*I, *Hin*dIII–*Sty*I, *Bam*HI–*Sty*I, *Pst*I–*Sty*I, and *Dra*I–*Sty*I fragments containing the nucleotides corresponding to bases −1206, −669, −582, −474, −270, −38 and −24 to +90 of the 5′ upstream and exon 1 sequence of the hOP gene were prepared from a pUC118 clone containing the 3.5 kb *Sac*I–*Sac*I fragment of the 5′ upstream, exon 1, intron 1, exon 2, intron 2, exon 3, and intron 3 of the hOP gene. These fragments were inserted into pUC119 vector. After propagation, each insert was cloned into pSVmCAT vector containing a multi-cloning site which was produced from the pSVOCAT vector. To construct 5′ deletion mutants, the fragments containing bases −474 to +90 and −270 to +90 were linearized with *Xba*I and *Kpn*I and with *Bam*HI and *Kpn*I respectively. The resulting DNAs were 5′ deleted using a kit (Takara) according to the manufacturer's instructions. Briefly, the DNAs were digested with exonuclease III. Aliquots (5 μl) were removed at 1 min intervals and mixed with 50 μl of 2 × Mung bean nuclease buffer. After heating for 5 min at 65 °C, the samples were incubated with Mung bean nuclease for 60 min at 37 °C. The 5′-deleted samples were then digested with Klenow fragment, followed by T4 DNA ligase treatment. To determine the extent of deletion into the 5′ end of the promoter, each mutant was sequenced as described previously [31]. Each 5′-deleted DNA was digested with *Sac*I and *Sph*I, and inserted into pSVmCAT. Plasmid DNA was prepared by alkaline lysis followed by two centrifugation steps through CsCl to isolate supercoiled plasmid DNA.

DNA transfections and CAT assay

Plasmid DNA was transfected into the human monocyte cell line SCC-3 and the human B lymphoma cell line Raji by electroporation. Cells (4×10^6) in 0.8 ml of PBS were transfected in the presence of 3 μg of relevant plasmids and 5 μg of DEAE–dextran [32]. Electroporation was at 300 V with a capacitance of 500 and 250 μF for SCC-3 and Raji cells respectively. After electroporation, the cells were suspended in 8 ml of DMEM supplemented with 10% fetal calf serum, seeded at 2×10^6 cells/plate, and cultured for 48 h. Cells were washed three times with PBS, and resuspended in 180 μl of 0.25 M Tris/HCl, pH 8.0. Extracts were prepared by freeze–thaw and then centrifuged. Supernatants were collected and assayed for protein according to the method of Bradford [33]. CAT was assayed as described previously [34]. The extract was incubated with [dichloroacetyl-1,2-^{14}C]chloramphenicol and 0.2 mg of acetyl-CoA for 1 h at 37 °C, and the products were separated by t.l.c. CAT activities were corrected for variations in transfection efficiencies, by co-transfecting the pCH110 containing the β-galactosidase gene and normalized with respect to those obtained with pSV2CAT.

Oligonucleotides

Complementary oligonucleotides were designed and synthesized on a Biosearch DNA synthesizer (Model 8700). The oligonucleotides (10 μg of each strand) were annealed in a solution containing 50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 10 mM spermidine and 1 mM EDTA, heated to 90 °C for 2 min, and cooled at 65 °C for 10 min, 37 °C for 10 min, and room temperature for 5 min.

Extract preparation

Nuclear extracts from SCC-3 and Raji cells were prepared as described by Dignam et al. [35]. Briefly, ~ $2–4 \times 10^8$ cells were harvested, pelleted, and lysed by 10 strokes of a Potter all-glass homogenizer in ice-cold buffer (two packed-cell pellet volumes) containing 10 mM Hepes (pH 7.9 at 4 °C), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT and antibiotics [0.5 mM phenylmethanesulphonyl fluoride (PMSF), 0.3 μg/ml leupeptin

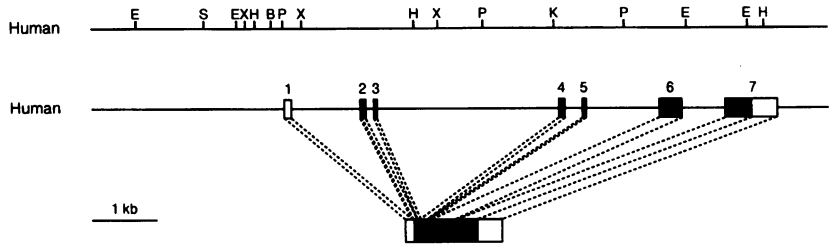

Figure 1 Exon and intron organization of the hOP gene

Exons are boxed, filled boxes are coding regions, and open boxes are untranslated regions. Abbreviations for restriction sites: E, *Eco*RI; S, *Sac*I; X, *Xba*I; H, *Hin*dIII; B, *Bam*HI; P, *Pst*I; and K, *Kpn*I.

```
-67 CTCCGCCTCCCTGTGTTGGTGGAGGATGTCTGCAGCAGCATTTAAATTCTGGGAGGGCTTGGTTGTCAGCAGCAGCAGGAGGAGGCAGAGCACAGCATCG Exon 1
+34 TCGGGACCAGACTCGTCTCAGGCCAGTTGCAGCCTTCTCAGCCAAACGCCGACCAAGGTACAGCTTCAGTTTGCTACTGGGTTGTGCATTCAGCTGAATT
    .....Intron 1 (1083 bp).....
1134 CTATGAAGATGTCAGCTATTCCTTATGAAATATTTTGCAGGAAAACTCACTACCATGAGAATTGCAGTGATTTGCTTTTGCCTCCTAGGCATCACCTGTG Exon 2
                                                    M  R  I  A  Y  I  C  F  C  L  L  G  I  T  C
                                                                  -16
1234 CCATACCAGTGAGTACAGTTGCATCTTAAAGAAAATTCCTGAAAATAACTGAATTGTGTGCTTCCATGTGCTAGGAGGACATTCTTGTAATCTTTCTTCA
     A  I  P
           2
    .....Intron 2 (109 bp).....
1334 TCTTTTCTGTTTCTAAGGTTAAACAGGCTGATTCTGGAAGTTCTGAGGAAAAGCAGGTAAGCATCTTTTATGTTTTTATATAGTTAATCATTTACTCAAT Exon 3
                       V  K  Q  A  D  S  G  S  S  E  E  K  Q
                          3                          15
    .....Intron 3 (2942 bp).....
4234 CTATATnCTTCCATCAAGACTAGTGAAGAATGGTTGTTTTTTCCATTCATCCCTACATTTCTTTTTTTAATAATGATAAACATGCAACTTTTTTGTAGCCT Exon 4
                                                                                                  L
                                                                                                 16
4334 TTACAACAAATACCCAGATGCTGTGGCCACATGGCTAAACCCTGACCCATCTCAGAAGCAGAATCTCCTAGCCCCACAGGTATTTTTAAACTTCTCATAA
       Y  N  K  Y  P  D  A  V  A  T  W  L  N  P  D  P  S  Q  K  Q  N  L  L  A  P  Q
                                                                                    42
    .....Intron 4 (265 bp).....
4634 TAGGCATGTGTGATGCGCACTAACACGTGCCATTCCTTCTTCAGAATGCTGTGTCCTCTGAAGAAACCAATGACTTTAAACAAGAGGTAAGTTCTCATTT Exon 5
                                                N  A  V  S  S  E  E  T  N  D  F  K  Q  E
                                               43                                   56
    .....Intron 5 (1039 bp).....
5734 TTTTCCCGGCCATCTTAATTTTCAGACCCTTCCAAGTAAGTCCAACGAAAGCCATGACCACATGGATGATATGGATGATGAAGATGATGATGACCATGTG Exon 6
                              T  L  P  S  K  S  N  E  S  H  D  H  M  D  D  M  D  D  E  D  D  D  D  H  V
                             57
5834 GACAGCCAGGACTCCATTGACTCGAACGACTCTGAGTATGTAGATGACACTGATGATTCTCACCAGTCTGATGAGTCTCACCATTCTGATGAATCTGATG
      D  S  Q  D  S  I  D  S  N  D  S  D  D  V  D  D  T  D  D  S  H  Q  S  D  E  S  H  H  S  D  E  S  D
5934 AACTGGTCACTGATTTTCCCACGGACCTGCCAGCAACCGAAGTTTTCACTCCAGTTGTCCCCACAGTAGACACATATGATGGCCGAGGTGATAGTGTGGT
     E  L  V  T  D  F  P  T  D  L  P  A  T  E  V  F  T  P  V  V  P  T  V  D  T  Y  D  G  R  G  D  S  V  V
6034 TTATGGACTGAGGTCAAAATCTAAGAAGTTTCGCAGACCTGACATCCAGGTAAATCCTTTAACAGACACACCTGATGGTTCTGACTAGCGCTCAAGTCTA
       Y  G  L  R  S  K  S  K  K  F  R  R  P  D  I  Q
                                                  164
    .....Intron 6 (692 bp).....
6734 AGCCGTTCATATAATTATTCTTCATTTGTGCCGTGATTCAGTACCCTGATGCTACAGACGAGGACATCACCTCACACATGGAAAGCGAGGAGTTGAATGG Exon 7
                                                  Y  P  D  A  T  D  E  D  I  T  S  H  M  E  S  E  E  L  N  G
                                                 165
6834 TGCATACAAGGCCATCCCCGTTGCCCAGGACCTGAACGCGCCTTCTGATTGGGACAGCCGTGGGAAGGACAGTTATGAAACGAGTCAGCTGGATGACCAG
       A  Y  K  A  I  P  V  A  Q  D  L  N  A  P  S  D  W  D  S  R  G  K  D  S  Y  E  T  S  Q  L  D  D  Q
6934 AGTGCTGAAACCCACAGCCACAAGCAGTCCAGATTATATAAGCGGAAAGCTAATGATGAGAGCAATGAGCATTCCGATGTGATTGATAGTCAGGAACTTT
      S  A  E  T  H  S  H  K  Q  S  R  L  Y  K  R  K  A  N  D  E  S  N  E  H  S  D  V  I  D  S  Q  E  L
7034 CCAAAGTCAGCCGTGAATTCCACAGCCATGAATTTCACAGCCATGAAGATATGCTGGTTGTAGACCCCAAAAGTAAGGAAGAAGATAAACACCTGAAATT
      S  K  V  S  R  E  F  H  S  H  E  F  H  S  H  E  D  M  L  V  V  D  P  K  S  K  E  E  D  K  H  L  K  F
7134 TCGTATTTCTCATGAATTAGATAGTGCATCTTCTGAGGTCAATTAAAAGGAGAAAAAATACAATTTCTCACTTTGCATTTAGTCAAAAGAAAAAATGCTT
       R  I  S  H  E  L  D  S  A  S  S  E  V  N  ***
                                                 298
7234 TATAGGAAAATGAAAGAGAACATGAAATGCTTCTTTCTCAGTTTATTGGTTGAATGTGTATCTATTTGAGTCTGGAAATAACTAATGTGTTTGATAATTA
7334 GTTTAGTTTGTGGCTTCATGGAAACTCCCTGTAAACTAAAAGCTTCAGGGTTATGTCTATGTTCATTCTATAGAAGAAATGCAAACTATCACTGTATTTT
7434 AATATTTGTTATTCTCTCATGAATAGAAATTTATGTAGAAGCAAACAAAATACTTTTACCCACTTAAAAAGAGAATATAACATTTTATGTCACTATAATC
7534 TTTTGTTTTTTAAGTTAGTGTATATTTTGTTGTGATTATCTTTTTGTGGTGTGAATAAATCTTTTATGTTGAATGTAATAAGAATTTGGTGGTGTCAATT
```

Figure 2 Nucleotide sequence of the hOP gene

Numbering is relative to the transcription initiation site. The amino acid sequence deduced by translation of exons is shown below the nucleotide sequence. Amino acids of the signal peptide are given negative numbers. Exons are indicated on the right-hand margin. Asterisks indicate the stop codon.

0.3 μg/ml antipain, 0.3 i.u./ml aprotinin, 0.5 mg/ml benzamidine, 0.1 μg/ml chymostatin and 0.7 μg/ml pepstatin]. The homogenate was checked microscopically for cell lysis and centrifuged for 15 min at 1500 *g* to pellet nuclei. The nuclear pellets were resuspended in low-salt buffer containing 20 mM Hepes, pH 7.9, 25 % (v/v) glycerol, 20 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT and antibiotics as above. After adding high-salt buffer (replacing 20 mM KCl with 1.2 M KCl) with stirring gently on ice, the nuclei were pelleted by centrifugation for 30 min at 25000 *g* at 4 °C. The supernatant was dialysed against 50 vol. of buffer containing 20 mM Hepes (pH 7.9), 20 % glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 0.3 μg/ml leupeptin and 0.3 μg/ml antipain overnight at 4 °C, then centrifuged for 20 min at 25000 *g* and stored in aliquots at −80 °C.

Gel-shift assay

Gel-shift assays were performed as described with minor modifications [36]. DNA probes consisted of synthetic oligo-

nucleotides labelled with T4 polynucleotide kinase and [γ-^{32}P] ATP. DNA probes were incubated with protein fractions at room temperature for 30 min in 50 μl of binding reaction containing 10 mM Tris, pH 7.5, 50 mM NaCl, 0.62 mM $ZnSO_4$, 0.5 mM EDTA, 1 mM DTT, 5% glycerol and 2 μg of poly(dI-dC) (Pharmacia LKB Biotechnology Inc.). Binding reactions contained ~ 2 ng (2×10^4 c.p.m.) of probe and 20 μg of nuclear extract. Binding reactions were terminated with 5 μl of 50 mM EDTA containing 0.05% (v/v) Bromophenol Blue, 0.05% (v/v) Xylene Cyanol and 5% (v/v) glycerol. Mixtures were then electrophoresed on 5% acrylamide gels in 6.7 mM Tris, pH 7.9, 3.3 mM sodium acetate and 1 mM EDTA at 4 °C for 70 min at 170 V with buffer recirculation. Gels were transferred to filter paper, dried, and autoradiographed.

RESULTS

Isolation of the genomic hOP gene and nucleotide sequence of OP genomic DNA

To isolate the genomic hOP clone, we screened the human liver genomic DNA library with the hOP cDNA as a probe. From 10^6 recombinant phages, one positive clone, GOP-2, insert size 13 kb, was isolated and subjected to the restriction enzyme mapping (Figure 1). The genomic sequence, determined as described in the Experimental section, is shown in Figure 2. The coding region is split into 7 exons (Figures 1 and 2) like that of the mouse [16,23,24]. The combined coding sequence of the gene is almost identical with the published hOP cDNA sequences, except for one nucleotide substitution (T for C) at position 817 [4].

The sizes of each exon and intron were roughly comparable with those for the mOP gene except for intron 3, which was ~ 2.7-fold longer than that of the mOP gene (Figure 3a). This difference is attributable to an insertion of ~ 1750 bp immediately before exon 4 in the hOP gene.

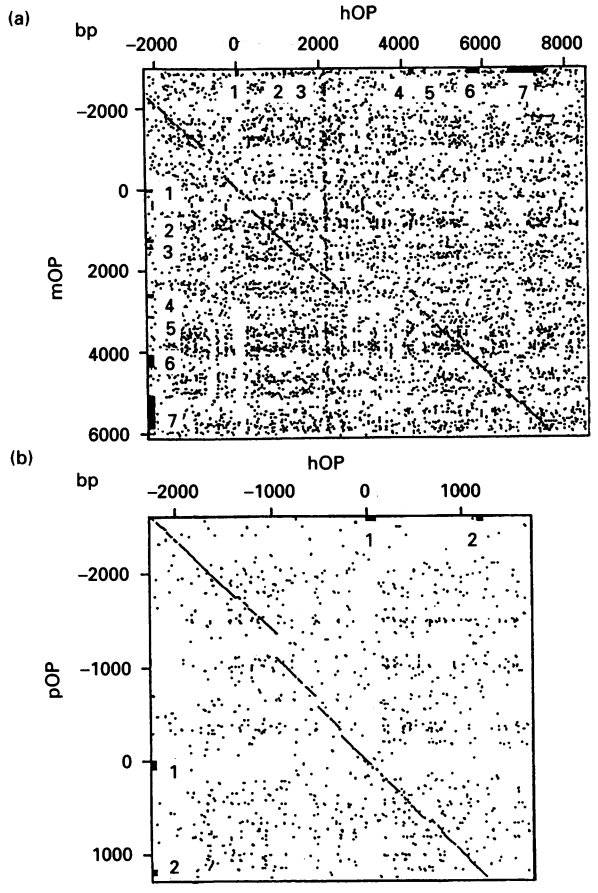

Figure 3 Dot-matrix analysis between the hOP, mOP and pOP genes

The plots were constructed using a computer program with standard parameters (match ≥ 14/21). The positions of the coding regions are indicated in filled boxes; a bold number shows exon number. (**a**) The hOP gene versus the mOP gene. (**b**) The hOP gene versus the pOP gene.

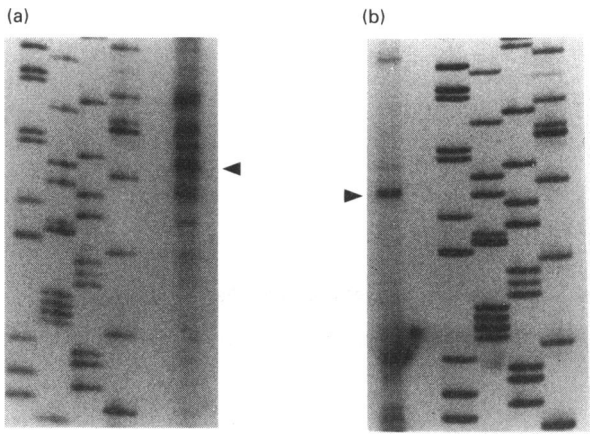

Figure 4 S1 nuclease mapping and primer extension analysis of hOP mRNA

(**a**) After digestion with S1 nuclease, as described in the Experimental section, the protected fragments were electrophoresed. Dideoxynucleotide sequencing reactions were electrophoresed in parallel as markers. The arrowhead indicates the site corresponding to that of the most predominant band in S1 nuclease mapping. (**b**) Total cellular RNA from THP-1 was annealed to a 30-base oligonucleotide specific for a 5′ region of hOP mRNA. The oligomers were extended and analysed as described in the Experimental section. Dideoxynucleotide sequencing reactions were electrophoresed in parallel as markers. The arrowhead indicates the 5′ transcriptional site (nucleotide +1; see Figure 2).

Identification of the 5′ end of hOP mRNA

The transcription initiation site of the mOP gene contains tandem AGC repeats [23], although initiation begins with the second AGC tandem repeat. The nucleotide sequence of the hOP gene also contains three AGC repeats in exon 1. Digestion with S1 nuclease resulted in major bands of 88, 90, 91, 93 and 97 nt protected fragments (Figure 4a). The strongest band, corresponding to the first nucleotide (A) of the triple repeats of the AGC at 90 nt, was subsequently designated mRNA residue +1 (Figure 2). To confirm that these S1-protected products indeed represented the 5′ end of the hOP mRNA, primer extension analysis was next performed. The primer extension product from this analysis revealed a major hOP fragment of 90 nt and several minor bands. This main band, indicated by an arrowhead in Figure 4(b), corresponded to the transcription initiation site determined by S1 nuclease protection.

Nucleotide sequence of the 5′ upstream of the hOP gene, comparison of the 5′ upstream and regulatory sequences with those of the mouse and porcine OP gene

The 5′ upstream sequence of the hOP gene and further upstream sequence of the mOP gene, part of which had previously been reported [23], were determined (the mouse sequence is not shown, but is available from GenBank/DDBJ, accession number

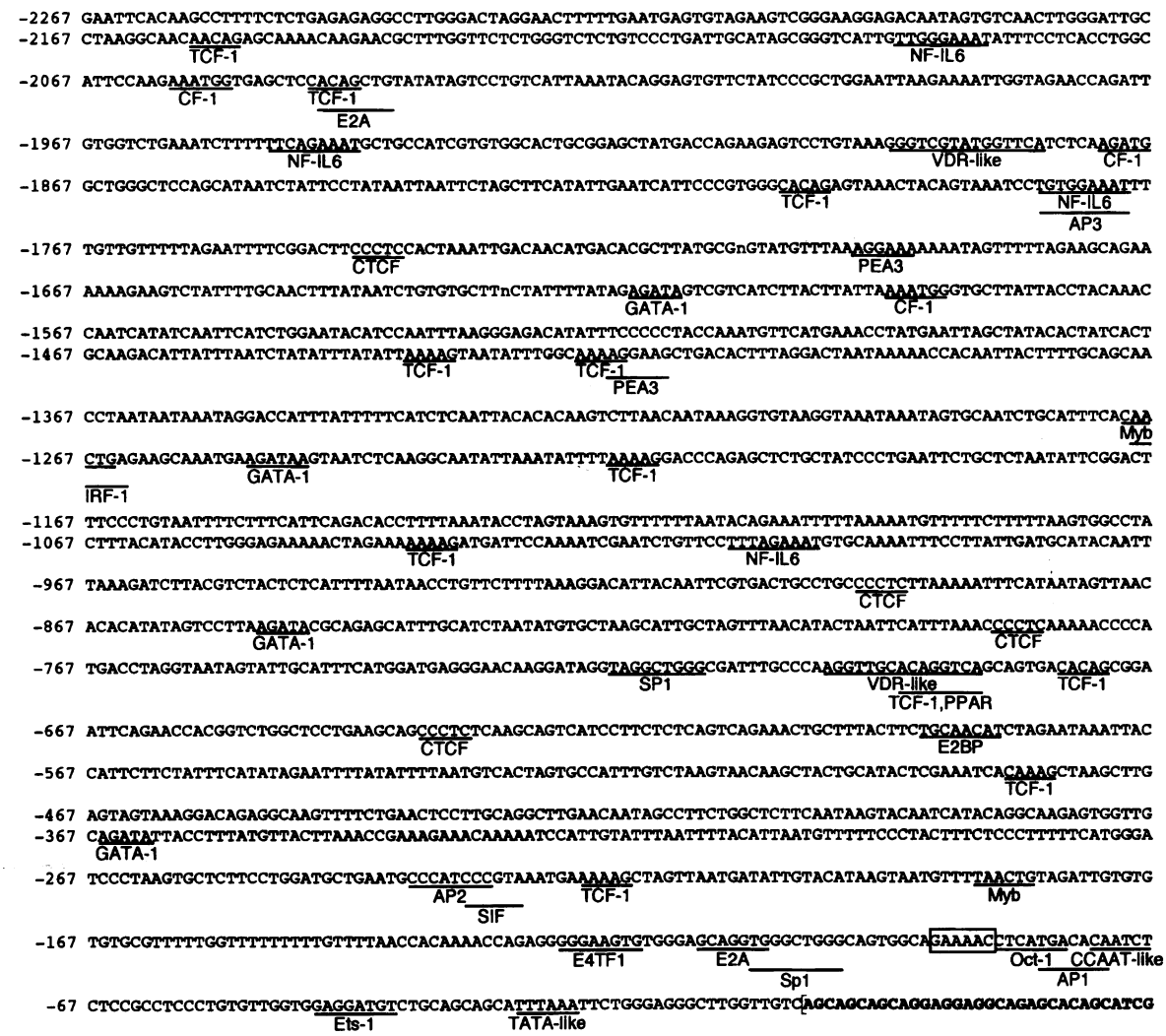

Figure 5 5′ upstream region of the hOP gene

Nucleotide positions are indicated at the left margin. The 5′ end of the cDNA is designated as nucleotide +1 of the gene. Regulatory motifs are indicated below the sequence; an open square is a *cis* region of hOP demonstrated by the present experiments.

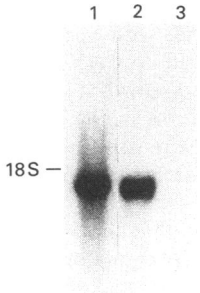

Figure 6 Expression of hOP mRNA in cell lines

Total cellular RNA was electrophoresed on a 1.5% (w/v) formaldehyde gel, blotted on to a nylon membrane, and hybridized with a ^{32}P-labelled hOP cDNA probe. Lanes contained RNA from the following sources: 1, SCC-3 cells; 2, THP-1 cells; 3, Raji cells.

D14816) and compared with each other and that of the porcine OP (pOP) gene [25]. As shown in dot-matrix comparison (Figure 3b), the 5′ upstream sequence of the hOP gene shows high similarity with that of the pOP gene.

The 5′ upstream sequence of the hOP gene contained a number of potential regulatory sequences (Figure 5) [37]. Many TATA- and CCAAT-like sequences were scattered over the 5′ upstream region. The vitamin-D-responsive sequence (VDR)-like motif [37], which was very similar to those in the mOP and pOP genes [23,25], was located at positions −1892 to −1878 and −698 to −684. GATA-1 was found at several positions, in which the sequence corresponding to that at position −851 to −847 was found in mouse and pig. The AP1 sequence, TGACACA [38], which was conserved in the mOP and pOP genes, was found at position −78 to −72. Other ubiquitous or tissue-specific motifs [37], including IRF-1, TCF-1, NF-IL6, CTCF, CF1, E2A, E2BP Myb, SP1, PPAR, AP2, SIF, E4TF1 [39], Oct-1 and Ets-1, were widely distributed. PEA3, a primary target of signal transduction for a variety of factors such as PMA, epidermal growth factor,

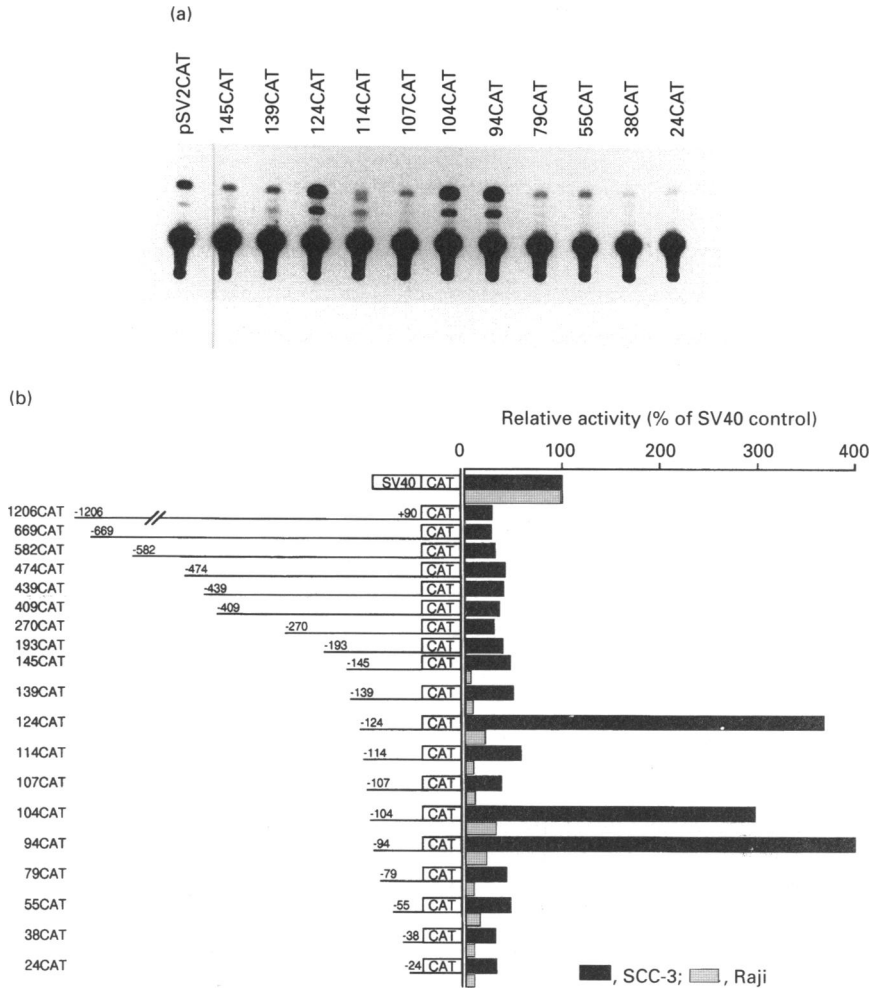

Figure 7 CAT analysis of the hOP gene

(**a**) Autoradiogram of CAT activity in extracts from SCC-3 cells transfected with various hOP–CAT fusion genes. A series of chimeric plasmids was constructed and transfected into SCC-3 cells. CAT activity expressed in a transient assay was measured. pSV2CAT contains simian virus 40 promoter only; 145CAT contains nucleotides −145 to +90; 139CAT contains nucleotides −139 to +90; 124CAT contains nucleotides −124 to +90; and so on. (**b**) Expression of 5′ deletion mutants of the hOP–CAT fusion gene in SCC-3 and Raji cells. A series of chimeric plasmids was constructed and transfected into SCC-3 and Raji cells: filled columns are CAT activities in SCC-3 cells; shaded columns are CAT activities in Raji cells. Nucleotide positions are indicated above the underline. The data represent the means of three independent transfections.

serum and oncogene products was found at positions −1695 to −1690, and −1418 to −1413.

CAT analysis of transcriptional regulatory regions upstream of the hOP gene

We examined expression of the hOP gene in SCC-3, THP-1 and Raji cells using the hOP cDNA as probe. SCC-3 and THP-1 expressed the hOP gene, whereas Raji did not (Figure 6). Next, we linked a series of upstream hOP gene fragments to the coding region of the CAT reporter gene to delineate the regions essential for transcription of the hOP gene (Figure 7). The 55 bp region was capable of directing a significant level of CAT synthesis in SCC-3 cells, although the region including 38 bp immediately upstream of the transcription start site showed no activity. The relative amount of reporter gene expression decreased by including additional 5′ sequences up to −79. Elongation of the 5′ upstream sequence up to −124, however, increased markedly the reporter gene expression in SCC-3 cells, although 107 CAT and 114 CAT (i.e. CAT sequences including positions up to −107 and −114) showed no CAT activity. A further stretch of the upstream sequence diminished the reporter gene expression, although marginal increase in the expression was observed by elongating the upstream sequence up to −474. Further elongation of the 5′ upstream sequence diminished reporter gene expression.

These results indicate that the 5′ flanking region at position −124 to −80 contains *cis* regulatory sequence specific for SCC-3 cells. Deletion of the sequence between −124 to −80 localized the *cis* region within the sequences at positions −124 to −115, −114 to −105, −104 to −95, and −94 to −80, although the sequence at position −104 to −95 might have no *cis* activity, because the activity of 104CAT does not exceed that of 94CAT (Figure 7). The sequence at position −107 to −105 could contain part of the negative *cis* region, because 107CAT abolished CAT activity.

We next tested CAT activities of 124CAT, 104CAT and 94CAT in HINS-B3 cells and found that these CAT clones

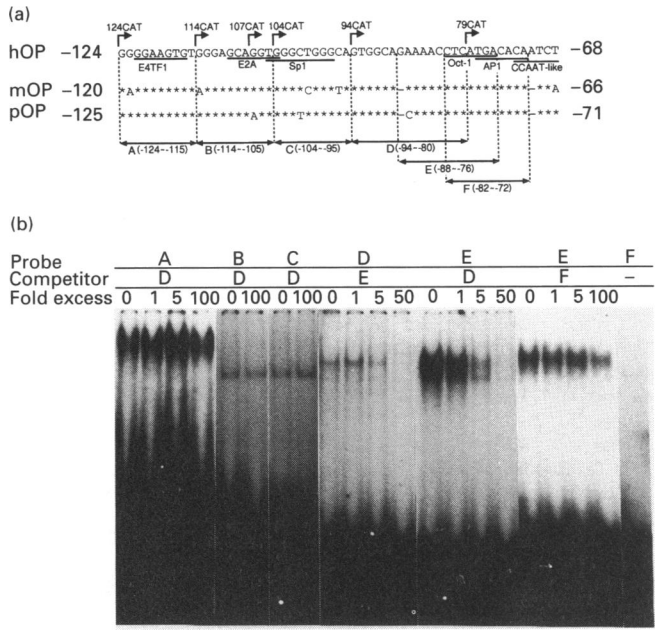

Figure 8 Gel-shift and competition assays

(**a**) The sequences conserved between the hOP, mOP and pOP genes in the 5′ region, and the synthetic oligonucleotides used in gel-shift and competition assays. Consensus E4TF1, E2A, Sp1, Oct-1, AP1 and CCAAT-like motifs are underlined. Asterisks indicate identities, and dashes indicate gaps. The letters A–F indicate the sequences used in gel-shift and competition assays; sequences are numbered relative to the transcription start sites of each gene. (**b**) Gel-shift and competition assays. ^{32}P-labelled probes A–F were incubated with crude nuclear extracts from SCC-3 cells in the absence or presence of unlabelled competitors.

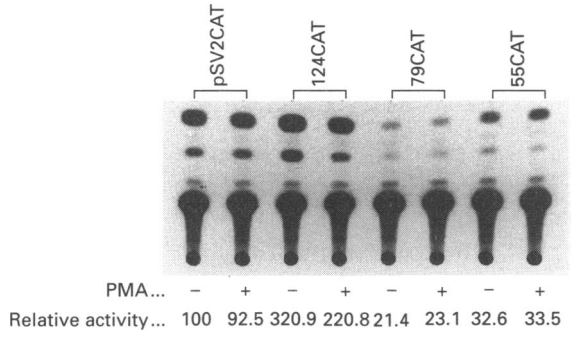

Figure 9 Effect of PMA on the expression of various hOP–CAT fusion genes in SCC-3 cells

A series of chimeric plasmids was transfected into SCC-3 cells. CAT activity in a transient assay was measured. pSV2CAT contains simian virus 40 promoter only; 124CAT contains nucleotides −124 to +90; 79CAT contains nucleotides −79 to +90; 55CAT contains nucleotides −55 to +90.

showed significant CAT activities, although the levels of the activities were low (results not shown).

Gel-shift analysis of protein(s) that specifically binds to the *cis* regulatory region

Gel-shift assays were performed using a series of synthetic oligonucleotides GATCGGGGAAGTGT, GATCGGGAGCAGGT, GATCGGGCTGGGCA, GATCGTGGCAGAAAACCTC, GATCGAAAACCTCATGA, and GATCCTCATGACACA, corresponding to positions −124 to −115, −114 to −105, −104 to −95, −94 to −80, −88 to −76, and −82 to −72, designated A, B, C, D, E and F respectively (Figure 8a). Each oligonucleotide formed a major band with nuclear extracts from SCC-3 cells. The band formed with probe A was competed out with corresponding unlabelled oligonucleotides (results not shown), but not unrelated ones (Figure 8b). The bands formed with probes B and C were identical and also competed out with corresponding unlabelled oligonucleotides (results not shown), but not unrelated ones (Figure 8b). The bands formed with probes D and E had identical electrophoretic mobility, although the band formed with probe D was weaker than that formed with probe E (Figure 8b). Probe F formed a very weak band. The major band formed with probe E was not competed with unlabelled F and A (results not shown). The *Bam*HI-site GATC introduced at the 5′ end of each probe had no effect on gel-shift analysis because no bands common to all probes were formed. These results suggested that the nucleotide position (G)AAAAC was involved in the formation of the major band. Each of the probes A–F formed a weak band with nuclear extracts from Raji cells which showed little or no CAT activity. The electrophoretic mobilities of all these bands, however, were different from those with nuclear extracts from SCC-3 cells (results not shown).

Effect of PMA on hOP promoter driven CAT expression

Finally, we compared CAT activities, using the 1206CAT, 582CAT, 439CAT, 270CAT, 124CAT, 79CAT and 55CAT clones in the presence of PMA (100 ng/ml). PMA, however, showed no effects in these CAT transfectants 3 and 16 h after stimulation. Some of the data are shown in Figure 9.

DISCUSSION

The nucleotide sequence of the hOP gene is considerably different from that of the mOP gene, as dot-matrix analysis revealed. The most extreme difference is that in intron 3 of hOP, which has a nucleotide insertion of ~ 1.8 kb. Considerable differences are found in exon 1 and intron 1 between the mouse and porcine genes. A region (~ 285 bp) immediately upstream of the hOP cap site was highly conserved. In contrast, a region further upstream of the hOP gene showed no similarity to that of the murine gene. The 5′ upstream sequence of the hOP gene, however, shows a significant similarity with that of the pOP gene. There are multiple potential consensus regulatory sequences in the OP genes. The numbers of consensus sequences shared by three species, however, are relatively few.

The hOP gene contains a TATA-like sequence, TTTAAA, at position −27 to −22, which had been reported in the mOP and pOP genes. CAT analysis, however, showed that 38CAT had no CAT activity in SCC-3 cells, suggesting that the TATA-like sequence is not functional, although further deletion of the 5′ sequence of 38CAT might reveal CAT activity. Instead, elongation of the 5′ upstream sequence up to −55 demonstrated slight CAT activity. This region contains an Ets-1-binding sequence.

Several 5′ upstream regions of the mOP gene have been reported to be responsible for constitutive and inducible expression [23,40], of which the region at position −543 to +79 shows the strongest activity when tested in a murine epidermal cell line JB6. Activities directed by these regions are enhanced by PMA [23]. Since the mOP gene is not normally expressed in JB6, the authors suggest that transcription from the mOP promoter is normally repressed or controlled at the post-transcriptional level. CAT analysis for the promoter region of the pOP gene revealed that a region at position −180 to +51 has markedly high promoter activity in calvarial bone cells [25]. CAT analysis of the

5′ upstream region of hOP revealed that regions at positions −124 to −115 and −94 to −80 accommodate strong *cis* regulatory sequences in SCC-3 cells, whereas the region at position −439 to −409 has a low level of *cis* activity. Nucleotide sequences of the major regions were highly conserved and showed *cis* activities in HINS-B3 cells, suggesting that the major *cis*-acting sequences are also important in other species.

Probe A formed one band, which may be the band related to E4TF1. Probe B formed one band which could be that formed with a negative regulatory factor. Probe C formed a band which was possibly the band associated with SP1, although the electrophoretic mobility of the band appeared identical with the band formed with probe B. However, the role of SP1 is uncertain because 104CAT did not confer additional CAT activity to 94CAT, and 107CAT abolished CAT activities. Thus, the region immediately downstream at position −107 to −105 is also likely to operate as a negative *cis* sequence. The bands formed with probe D were competed out with competitor E, and vice versa. These bands were not competed out with competitor F. Therefore, these results suggest that (G)AAAAC is important as a *cis* sequence. Since the signal intensities of the bands formed with F were less marked, involvement of Oct-1 may be minor or even non-existent. Taken together, these data suggest that (G)AAAAC could play a major role in the promoter activity in the hOP gene in SCC-3 cells. The region at position −124 to −115 containing the E4TF1 site may act to suppress the negative regulatory region between −107 and −105.

The experiments described herein have demonstrated that the conserved AP1 site at position −78 to −72 is not responsible for enhancement of CAT activity in SCC-3, although it is bound by a nuclear protein. 124CAT and 55CAT also showed no PMA inducibilities. CAT clones containing further upstream elements had no CAT activity (results not shown). VDR-like motifs are found at positions −698 to −684 and −1892 to −1878, and have been reported to be functional in mOP and pOP respectively [25,40]. Whether these sequences are functional in the hOP gene remains to be clarified.

This work was supported by grants from the Mitsuiseimei Kosei Jigyodan, and Ministry of Education, Science and Culture (no. 05670201).

REFERENCES

1 Oldberg, A., Franzen, A. and Heinegard, D. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 8819–8823
2 Miyazaki, Y., Setoguchi, M., Yoshida, S., Higuchi, Y., Akizuki, S. and Yamamoto, S. (1989) Nucleic Acids Res. **17**, 3298
3 Craig, A. M., Smith, J. H. and Denhardt, D. T. (1989) J. Biol. Chem. **264**, 9682–9689
4 Kiefer, M. C., Bauer, D. M. and Barr, P. J. (1989) Nucleic Acids Res. **17**, 3306
5 Wrana, J. L., Zhang, Q. and Sodek, J. (1989) Nucleic Acids Res. **17**, 10119
6 Castagnola, P., Bet, P., Quarto, R., Gennari, M. and Cancedda, R. (1991) J. Biol. Chem. **266**, 9944–9949
7 Reinholt, F. P., Hultenby, K., Oldberg, A. and Heinengard, D. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 4473–4475
8 Kasugai, S., Todescan, R., Nagata, T., Yao, K.-L., Butler, W. T. and Sodek, J. (1991) J. Cell. Physiol. **147**, 111–120
9 Kasugai, S., Nagata, T. and Sodek, J. (1992) J. Cell. Physiol. **152**, 467–477
10 Prince, C. W., Oosawa, T., Butler, W. T., Tomana, M., Bhown, A. S., Bhown, M. and Schrohenloher, R. E. (1987) J. Biol. Chem. **262**, 2900–2907
11 Somerman, M. J., Prince, C. W., Sauk, J. J., Foster, R. A. and Butler, W. T. (1987) J. Bone Miner. Res. **2**, 259–265
12 Patarca, R., Freeman, G. J., Singh, R. P., Wei, F.-Y., Durfee, W. T., Blattner, F., Regnier, D. C., Kozak, C. A., Mock, B. A., Morse, H. C., III, Jerrells, T. R. and Cantor, H. (1989) J. Exp. Med. **170**, 145–161
13 Yoon, K., Buenaga, R. and Rodan, G. A. (1987) Biochem. Biophys. Res. Commun. **13**, 1129–1136
14 Mark, M. P., Prince, C. W., Gay, S., Austin, R. L. and Butler, W. T. (1988) Cell Tissue Res. **251**, 23–30
15 Nomura, S., Willis, A. J., Edwards, D. R., Heath, J. K. and Hogan, B. L. M. (1988) J. Cell Biol. **106**, 441–450
16 Miyazaki, Y., Setoguchi, M., Yoshida, S., Higuchi, Y., Akizuki, S. and Yamamoto, S. (1990) J. Biol. Chem. **265**, 14432–14438
17 Rodan, S. B., Wesolowski, G., Yoon, K. and Rodan, G. A. (1989) J. Biol. Chem. **264**, 19934–19941
18 Noda, M., Yoon, K., Prince, C. W., Butler, W. T. and Rodan, G. A. (1988) J. Biol. Chem. **263**, 13916–13921
19 Noda, M., Vogel, R. L., Hasson, D. M. and Rodan, G. A. (1990) Endocrinology (Baltimore) **127**, 185–190
20 Noda, M. and Rodan, G. A. (1989) J. Cell Biol. **108**, 713–718
21 Smith, J. H. and Denhardt, D. T. (1987) J. Cell. Biochem. **34**, 13–22
22 Senger, D. R., Perruzzi, C. A., Gracey, C. F., Papadopoulos, A. and Tenen, D. G. (1988) Cancer Res. **48**, 5770–5774
23 Craig, A. M. and Denhardt, D. T. (1991) Gene **100**, 163–171
24 Behrend, E. I., Chambers, A. F., Wilson, S. M. and Denhardt, D. T. (1993) J. Biol. Chem. **268**, 11172–11175
25 Zhang, Q., Wrana, J. L. and Sodek, J. (1992) Eur. J. Biochem. **207**, 649–659
26 Kimura, Y., Toki, H., Okabe, K., Suzuki, K., Hayakawa, T. and Miyamoto, K. (1986) Jpn. J. Cancer Res. **77**, 862–865
27 Tsuchiya, S., Yamabe, M., Yamaguchi, Y., Kobayashi, Y., Konno, T. and Tada, K. (1980) Int. J. Cancer **26**, 171–176
28 Pulvertaft, R. J. V. (1964) Lancet **1**, 238–239
29 Higuchi, Y., Setoguchi, M., Yoshida, S., Akizuki, S. and Yamamoto, S. (1988) Oncogene **2**, 515–521
30 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
31 Berk, A. J. and Sharp, P. A. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 1274–1278
32 Gauss, G. H. and Lieber, M. R. (1992) Nucleic Acids Res. **20**, 6739–6740
33 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
34 Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982) Mol. Cell. Biol. **2**, 1044–1051
35 Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) Nucleic Acids Res. **11**, 1475–1489
36 Matsuura, K., Ishida, T., Setoguchi, M., Higuchi, Y., Akizuki, S. and Yamamoto, S. (1992) J. Biol. Chem. **267**, 21787–21794
37 Faisst, S. and Meyer, S. (1992) Nucleic Acids Res. **20**, 3–26
38 Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P. and Karin, M. (1987) Cell **49**, 729–739
39 Watanabe, H., Wada, T. and Handa, H. (1990) EMBO J. **9**, 841–847
40 Noda, M., Vogel, R. L., Craig, A. M., Prahl, J., DeLuca, H. F. and Denhardt, D. T. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 9995–9999

Received 8 February 1994/18 April 1994; accepted 6 May 1994